

**ΕΝΩΣΗ ΕΛΛΗΝΩΝ ΦΥΣΙΚΩΝ
11^Ο ΠΑΝΕΛΛΗΝΙΟ ΣΥΝΕΔΡΙΟ ΦΥΣΙΚΗΣ**

**«ΟΙ ΝΕΟΙ ΟΡΙΖΟΝΤΕΣ ΤΗΣ ΦΥΣΙΚΗΣ ΕΠΙΣΤΗΜΗΣ ΣΤΟΝ ΑΙΩΝΑ ΜΑΣ
ΣΤΗΝ ΕΡΕΥΝΑ, ΣΤΗΝ ΤΕΧΝΟΛΟΓΙΑ ΚΑΙ ΣΤΗΝ ΕΚΠΑΙΔΕΥΣΗ»**

30-31 ΜΑΡΤΙΟΥ, 1-2 ΑΠΡΙΛΙΟΥ 2006-02-25 ΛΑΡΙΣΑ

«Ας ακούσουμε μια εικόνα»

Χαράλαμπος Χ. Σπυρίδης¹ και Άγγελος Κ. Μουστάκας²

¹ Καθηγητής Μουσικής Ακουστικής, Πληροφορικής –Διευθυντής Εργαστηρίου Μουσικής Ακουστικής τεχνολογίας, Τμήματος Μουσικών Σπουδών Πανεπιστημίου Αθηνών.

² Φυσικός.

1. Περίληψη

Αντικείμενο της εισηγήσεως αποτελεί η παρουσίαση λογισμικού, που συγγράψαμε, δια του οποίου δημιουργείται το ηχητικό ανάλογο μιας εικόνας. Το λογισμικό δι' ενός αλγορίθμου παράγει μουσικές μελωδίες μετατρέποντας την RGB πληροφορία για το χρώμα κάθε σημείου μιας εικόνας ψηφιακής μορφής, σε ηχητικές συχνότητες της συγκεκριμένης ευρωπαϊκής κλίμακας.

Πιο συγκεκριμένα, μετά την ανάγνωση της RGB χρωματικής πληροφορίας για κάθε σημείο μιας εικόνας αποθηκευμένης σε υπολογιστή,

- μετατρέπουμε τη χρωματική πληροφορία σε μονοχρωματική συχνότητα του ορατού φάσματος και παράγουμε νέα εικόνα.
- μετασχηματίζουμε τις συχνότητες του ορατού φάσματος σε συχνότητες του ακουστού φάσματος.
- μέσω μιας μαθηματικής διαδικασίας ομαδοποιούμε τις ακουστές συχνότητες και τις ανάγουμε σε φθόγγους της συγκεκριμένης ευρωπαϊκής κλίμακας, συγκεκριμένου μουσικού ύψους και διάρκειας.

Μας δίδεται, έτσι, η δυνατότητα να ακούμε την εικόνα μέσω του μουσικού αναλόγου της, που προέκυψε.

2. Προλεγόμενα

Κυρίες και Κύριοι σύνεδροι, επιθυμών να σας προϊδεάσω για την επομένη εισήγηση με θέμα «Ας ακούσουμε μια εικόνα», θα σας αναφέρω δύο αποφθέγματα. Το πρώτο μας το διασώζει ο Πολύβιος στο έργο του *Ιστορίαι* και το δεύτερο ο Σέξτος ο Εμπειρικός στο έργο του *Προς Μαθηματικούς*. Αμφότερα αποδίδονται στον μεγάλο προσωκρατικό φιλόσοφο Ηράκλειτο τον Εφέσιο (540-480 π.Χ.).

ὄφθαλμοὶ γὰρ τῶν ὄτων ἀκριβέστεροι μάρτυρες
Πολυβίου, *Ιστορίαι*, 12, 27, 1, 5

[τα μάτια είναι ακριβέστεροι μάρτυρες απ' ό,τι τα αυτιά]

Η κοινωνική μας δραστηριότητα καθίσταται δυνατή δια της αισθήσεως της οράσεως και της ακοής. Τα μάτια, τα όργανα της οράσεως, θεωρούνται πιστότεροι μάρτυρες από τα αυτιά, τα όργανα της ακοής.

Στο απόσπασμα αυτό ο Ηράκλειτος διαχωρίζει τις πηγές της γνώσεως λέγοντας ότι τα μάτια είναι ακριβέστεροι μάρτυρες, όσον αφορά στην ουσία των όντων, διότι από τη μορφή και το είδος συλλαμβάνομε την εντελέχεια την οποία πραγματοποιούν τα όντα κατά την εδώ πορεία τους.

κακοὶ μάρτυρες ἀνθρώποισιν ὀφθαλμοὶ καὶ ὄτα
βαρβάρους ψυχὰς ἔχόντων
Σέξτος Εμπειρικός, *Προς Μαθηματικούς*, 7, 126, 7-8

[Τα μάτια και τα αυτιά είναι κακοί μάρτυρες για τους ανθρώπους που έχουν βάρβαρη (=ακαλλιέργητη) ψυχή]

Κατά τον Ηράκλειτο τα μάτια και τα αυτιά μαρτυρούν για την ύπαρξη των όντων. Η γνώση, όμως, επιτυγχάνεται μόνον όταν ο άνθρωπος αντιλαμβάνεται σωστά τα όσα του προσφέρουν οι αισθήσεις. Άρα ο άνθρωπος αξιολογείται ανάλογα με την ικανότητά του να επεξεργάζεται τις πληροφορίες δια των αισθήσεων με σκοπό να επιτύχει την κατανόηση του όντος.

Βάρβαρος (=ακαλλιέργητος) είναι ο άνθρωπος που δεν μεθίσταται των φαινομένων, ώστε να προσεγγίσει την αλήθεια.

Οι πρόγονοί μας, κυρίες και κύριοι σύνεδροι, δια μέσου της φιλοσοφίας είχαν συνείδηση ότι ο πολιτισμός χαρακτηρίζεται από την επεξεργασία της αισθήσεως.

Αφού, λοιπόν, τα μάτια είναι ακριβέστεροι μάρτυρες απ' ό,τι είναι τα αυτιά, μια εικόνα, που ευχαριστεί τα μάτια, θα μπορούσε – με κάποιο τρόπο – να ευχαριστήσει και τα αυτιά;

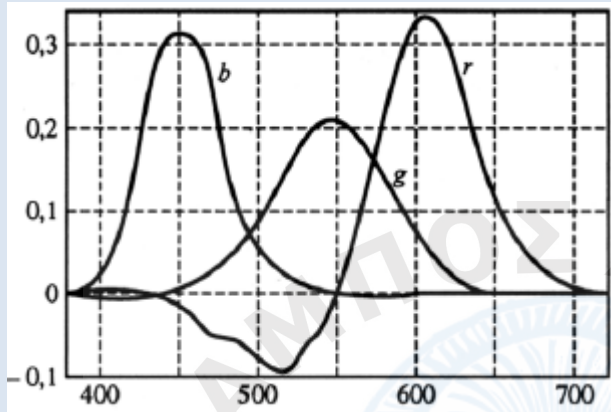
Αυτό το ερώτημα μας ωδήγησε στη συγκεκριμένη έρευνα από την οποία προσπαθήσαμε να βρούμε την απάντησή του. Η απάντηση εξαρτάται από τον χαρακτηρισμό βαρβάρους ψυχάς, που θα μπορούσε να εκλιφθεί ως ψυχές εθισμένες σε εικόνες συγκεκριμένης τεχνοτροπίας και σε ακούσματα συγκεκριμένης μουσικής νοοτροπίας. Προβληματισθείτε βλέποντας και ακούοντας τα της εισηγήσεως και καταλήξτε στη δική σας απάντηση.

3. Απόδοση χρωμάτων και μοντέλο RGB.

Το χρώμα στους υπολογιστές παριστάνεται ως μείξη τριών βασικών χρωμάτων. Πιο συγκεκριμένα χρησιμοποιείται το χρωματικό μοντέλο RGB και κάθε χρώμα προκύπτει ως μίξη του Κόκκινου (R), του Πράσινου (G) και του Μπλε (B). Για κάθε σημείο μιας εικόνας (pixel), όπως αυτή παριστάνεται σε οθόνη πραγματικού χρώματος, αφιερώνονται 3 bytes πληροφορίας. Το καθένα από αυτά, με τιμή από 0 ως 255 ($2^8 = 256$), αφορά στο «ποσοστό» συμμετοχής καθενός από τα τρία βασικά χρώματα στην σύνθεση του αποδιδόμενου χρώματος. Οι καμπύλες μείξης RGB για την σύνθεση των λοιπών χρωμάτων παριστάνονται στο σχήμα 1.

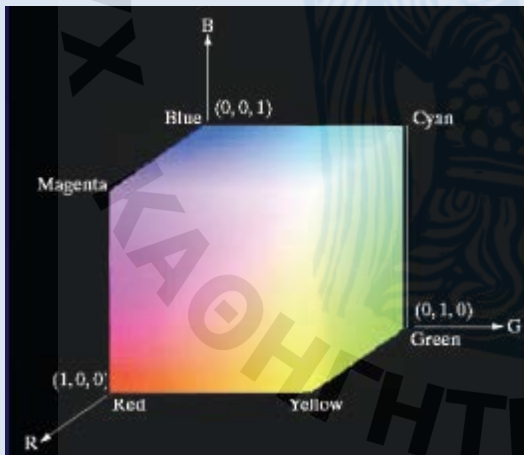
Εξετάζοντας τις καμπύλες μείξης των βασικών χρωμάτων του μοντέλου RGB γίνεται φανερή η αδυναμία απόδοσης κάποιων συχνοτήτων του ορατού φάσματος, διότι, ενώ

αυτό λειτουργεί προσθετικά, σε κάποια περιοχή μηκών κύματος του ορατού φάσματος η καμπύλη του ερυθρού χρώματος λαμβάνει αρνητικές τιμές. Η αδυναμία αποδόσεως όλων των χρωμάτων από το μοντέλο RGB οφείλεται στο γεγονός ότι η παράσταση των χρωμάτων μεταφέρεται από το πεδίο των μηκών κύματος στο πεδίο απεικονίσεως¹.



Σχήμα 1: Οι καμπύλες μείξης RGB για την σύνθεση των λοιπών χρωμάτων.

Πιο συγκεκριμένα το μοντέλο RGB ορίζει έναν τρισδιάστατο χώρο, ο οποίος περιλαμβάνει όλα τα χρώματα που είναι δυνατόν να παρασταθούν από έναν υπολογιστή.



Σχήμα 2: Ο τρισδιάστατος χώρος που ορίζει το RGB μοντέλο, ο οποίος περιλαμβάνει όλα τα χρώματα που είναι δυνατόν να παρασταθούν από έναν υπολογιστή.

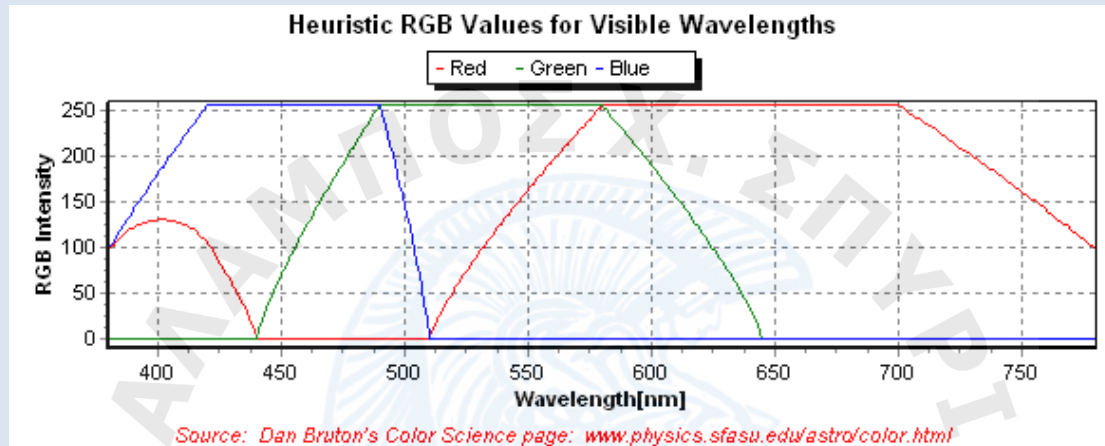
Κάθε άξονας του τρισδιάστατου αυτού χώρου λαμβάνει ($2^8 = 256$) τιμές στο διάστημα $[0,255]$ ή $[0,1]$ σε κανονικοποιημένες τιμές. Το $(0,0,0)$ παριστάνει το μαύρο χρώμα και το $(1,1,1)$ το λευκό. Κάθε άλλο χρώμα προκύπτει ως γραμμική σύνθεση:

$$F = r R + g G + b B$$

¹ GARRETT M. JOHNSON & MARK D. FAIRCHILD, "Computer Synthesis of Spectroradiometric Images for Color Imaging Systems Analysis", The Sixth Color Imaging Conference: Color Science, Systems, and Applications

Εξ αιτίας αυτής της ιδιομορφίας της χρωματικής πληροφορίας στους υπολογιστές, δεν υπάρχει μια μονοσήμαντη αντιστοίχιση συχνοτήτων και RGB τιμών.

Αναγκαστήκαμε, λοιπόν, να χρησιμοποιήσουμε κάποια εвриματική μέθοδο. Επιλέξαμε τον αλγόριθμο Bruton² λόγω του πολύ καλού αισθητικού αποτελέσματος με την έννοια ότι αποδίδει πολύ ικανοποιητικά τα χρώματα του ορατού φάσματος. Στο σχήμα 3 απεικονίζονται οι RGB τιμές για τα ορατά μήκη κύματος, όπως αυτά προκύπτουν από τον αλγόριθμο Bruton.



Σχήμα 3: Οι RGB τιμές για τα ορατά μήκη κύματος, όπως προκύπτουν από τον αλγόριθμο Bruton.

Ο αναφερθείς αλγόριθμος του Bruton μετατρέπει συχνότητες του ορατού φάσματος σε RGB τιμές. Δεν υπάρχει αντίστροφος αλγόριθμος, ο οποίος να μετατρέπει RGB τιμές σε μονοχρωματικές συχνότητες του ορατού φάσματος.

Προκειμένου, όμως, να έχουμε τιμές συχνοτήτων του ορατού φάσματος για τις RGB τιμές του χρώματος (r,g,b) εκάστου pixel της εικόνας, που θα μετατρέψουμε σε μουσική, εφαρμόσαμε την ακόλουθη διαδικασία:

Κάθε τιμή χρώματος (r,g,b) εκάστου pixel της εικόνας, που θα μετατρέψουμε σε μουσική, πρέπει να αντιστοιχηθεί σε κάποια τιμή χρώματος (r',g',b') του ορατού φάσματος, η οποία προέκυψε από τον αλγόριθμο του Bruton.

Η αντιστοίχιση επιτυγχάνεται ελαχιστοποιώντας τη συνάρτηση

$$f(r',g',b') = \sqrt{(r-r')^2 + (g-g')^2 + (b-b')^2},$$

η οποία παριστάνει την μεταξύ τους απόσταση στον τρισδιάστατο RGB χώρο.

Για προφανείς πρακτικούς λόγους μνήμης και διάρκειας της μουσικής σύνθεσης, που θα προκύψει, η ανωτέρω διαδικασία δεν εφαρμόζεται σε όλα τα σημεία της εικόνας, αλλά στα σημεία εκείνα, που επιλέγονται με βάση κάποιο καθοριζόμενο από τον χρήστη (user) βήμα κατά την οριζόντια και κατά την κατακόρυφο διεύθυνση.

² DAN BRUTON, "Color Science" Web Page, <http://www.physics.sfasu.edu/astro/color.htm>

Υλοποιώντας, λοιπόν, τις αναφερθείσες ενέργειες, που φαίνονται στο σχήμα 4, μετατρέπουμε τις RGB τιμές για το χρώμα κάθε σημείου (pixel) μιας εικόνας σε συχνότητες του ορατού φάσματος.



Σχήμα 4: Οι ενέργειες δια των οποίων μετατρέπουμε τις RGB τιμές για το χρώμα κάθε σημείου (pixel) μιας εικόνας σε συχνότητες του ορατού φάσματος.

Επίσης, έχοντας νέες RGB τιμές για το χρώμα του κάθε σημείου (που αντιστοιχούν σε συχνότητες του ορατού φάσματος) ξανασχεδιάζουμε την εικόνα χρησιμοποιώντας αυτές τις τιμές.

4. Μετασηματισμός Οπτικών συχνοτήτων σε Ακουστές συχνοτήτες

Εν συνεχεία, μετατρέπουμε τις συχνότητες του οπτικού φάσματος σε συχνότητες του ακουστού φάσματος χρησιμοποιώντας την παρακάτω σχέση μετασηματισμού:

$$\frac{\log(f_{\text{οπτικό}}) - \log(f_{\text{οπτικό}}^{\min})}{\log(f_{\text{οπτικό}}^{\max}) - \log(f_{\text{οπτικό}}^{\min})} = \frac{\log(f_{\text{ακουστό}}) - \log(f_{\text{ακουστό}}^{\min})}{\log(f_{\text{ακουστό}}^{\max}) - \log(f_{\text{ακουστό}}^{\min})}$$

ως συνέπεια του ψυχοφυσικού νόμου των Weber-Fechner ότι δηλαδή το αίσθημα είναι ανάλογο του λογαρίθμου του ερεθίσματος. Με άλλα λόγια αυτό που βλέπουμε και αυτό που ακούμε είναι ανάλογο των λογαρίθμων των συχνοτήτων των κυματικών (οπτικών και ακουστικών, αντιστοίχως) ερεθισμάτων.

Τα όρια για το ακουστό φάσμα –υπό την έννοια της ανθρώπινης μουσικής δραστηριότητας- είναι επιλέξιμα εντός του συχνοτικού εύρους οκτώ μουσικών οκτάβων και συγκεκριμένα από 16,352 Hz έως 7.902,1 Hz (C0-B8). Τα όρια του ορατού φάσματος είναι προφανώς τα μήκη κύματος από 380nm ως 780nm.

5. Ομαδοποίηση τιμών με χρήση του Συντελεστή Συσχέτισης

Έχοντας ως στόχο να καταλήξουμε σε μουσική μελωδία, πρέπει με κάποιο τρόπο να ομαδοποιήσουμε τις προκύπτουσες συχνότητες έτσι, ώστε να διαμορφωθούν φθόγγοι συγκεκριμένου μουσικού ύψους και χρονικής διάρκειας.

Επιλέχθηκε να χρησιμοποιήσουμε τη διαδικασία με τον Συντελεστή Συσχέτισης r που περιγράφουν στην εργασία τους οι Σπυρίδης, Ρουμελιώτης, Παπαδημητράκη-Χλίχλια³. Ο Συντελεστής Συσχέτισης

$$r = \frac{n \sum xy - \sum x \sum y}{\sqrt{[n \sum x^2 - (\sum x)^2][n \sum y^2 - (\sum y)^2]}}$$

ως γνωστόν, περιγράφει την ποιότητα της προσαρμογής n πλήθους σημείων επ' ευθείας.

Εφ' όσον ο συντελεστής συσχέτισης ενός πλήθους n διαδοχικών συχνοτικών τιμών εμπίπτει εντός των ορίων, που καθορίζονται από τον χρήστη (user), τότε οι συχνότητες αυτές ομαδοποιούμενες δίδουν φθόγγο μουσικού ύψους ίσο με τη μέση τιμή των εν λόγω συχνοτήτων και διάρκειας ίσης προς το πλήθος των ομαδοποιούμενων τιμών.

6. Μετατροπή Ακουστών συχνοτήτων σε συχνότητες της συγκερασμένης ευρωπαϊκής μουσικής κλίμακας και δημιουργία πολυφωνικής μουσικής.

Το επόμενο βήμα είναι η αναγωγή των συχνοτήτων των φθόγγων, που προέκυψαν, σε συχνότητες της συγκερασμένης ευρωπαϊκής κλίμακας. Αυτό επιτυγχάνεται με την αντικατάσταση της συχνότητας εκάστου φθόγγου με την πλησιέστερή της συχνότητα της συγκερασμένης ευρωπαϊκής κλίμακας. Κατ' αυτόν τον τρόπο προκύπτει μελωδία με φθόγγους της ευρωπαϊκής μουσικής.

Τους φθόγγους αυτούς τους χωρίζουμε στον επιθυμητό από τον χρήστη αριθμό συνηχουσών φωνών ως εξής: η πρώτη γραμμή σημείων (pixel) της εικόνας αποδίδεται στην πρώτη φωνή, η δεύτερη γραμμή αποδίδεται στη δεύτερη φωνή κ.ο.κ μέχρις εξαντλήσεως του πλήθους των επιθυμητών συνηχουσών φωνών.

Η επόμενη γραμμή της εικόνας αποδίδεται και πάλι στην πρώτη φωνή κ.ο.κ. Δηλαδή οι n το πλήθος φωνές σαρώνουν ταυτόχρονα ζώνη εύρους n γραμμών σημείων της εικόνας.

Η μελωδία που παράγεται αποθηκεύεται σε αρχείο σε μορφή *midi* και είναι δυνατό να την ακούσουμε ή να την επεξεργαστούμε με οποιοδήποτε πρόγραμμα επεξεργασίας παρτιτούρας.

³ H. SPYRIDIS, E. ROUMELIOTIS, and H. PAPADIMITRAKI - CHLICHLIA, *A computer approach to the construction and analysis of a pitch-curve in music*, ACUSTICA, vol. 51, No 3, pp 180-182, (1982).

Το λογισμικό μας παρέχει τη δυνατότητα:

- επιλογής διαφορετικών μουσικών οργάνων για την απόδοση κάθε φωνής (μελωδίας)
- επιλογής του μέτρου,
- ρύθμισης του tempo,
- «εκτέλεσης» της πολυφωνικής μελωδίας και
- προβολής σε πίνακα των νοτών όλων των φωνών.

7. Αντιστροφή της διαδικασίας και επανασχεδίαση της αρχικής εικόνας

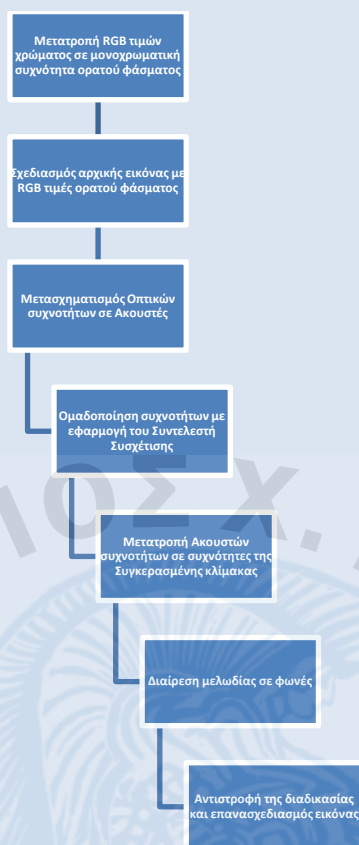
Από τη μελωδία που έχει προκύψει έχουμε μια τιμή συγκερασμένης συχνότητας του ακουστού φάσματος για κάθε σημείο (pixel) της αρχικής εικόνας. Με εφαρμογή της σχέσης μετασχηματισμού

$$\frac{\log(f_{\text{οπτικό}}) - \log(f_{\text{οπτικό}}^{\min})}{\log(f_{\text{οπτικό}}^{\max}) - \log(f_{\text{οπτικό}}^{\min})} = \frac{\log(f_{\text{ακουστό}}) - \log(f_{\text{ακουστό}}^{\min})}{\log(f_{\text{ακουστό}}^{\max}) - \log(f_{\text{ακουστό}}^{\min})}$$

για κάθε δεδομένη ακουστή συχνότητα προκύπτει η αντίστοιχή της συχνότητα του ορατού φάσματος. Τα όρια είναι για τις οπτικές συχνότητες εκείνα του ορατού φάσματος και για τις ακουστές εκείνα που χρησιμοποιήθηκαν κατά τον μετασχηματισμό των οπτικών συχνοτήτων σε ακουστές. Με την βοήθεια του αλγορίθμου του Bruton λαμβάνονται οι αντίστοιχες RGB τιμές και επανασχεδιάζεται η αρχική εικόνα έχοντας ως «πηγή» την μελωδία. Η προκύπτουσα εικόνα είναι κατά το μάλλον ή ήττον αλλοιωμένη, λόγω του ότι οι ακουστές συχνότητες αφενός είχαν αναχθεί –κατά τις απαιτήσεις της ευρωπαϊκής μουσικής- σε συχνότητες της συγκερασμένης ευρωπαϊκής κλίμακας και αφετέρου είχαν αλλοιωθεί λόγω της ομαδοποίησής τους με χρήση του συντελεστή συσχέτισης.

8. Ανακεφαλαίωση

Ανακεφαλαιώνοντας, οι διαδικασίες τις οποίες εκτελεί το λογισμικό που συγγράψαμε, συνοψίζονται στο βήματα του διαγράμματος στο σχήμα 5.



Σχήμα 5: Το διάγραμμα ροής του λογισμικού μας.